

Calidad de Información: una nueva herramienta para la investigación

Prof. Mag. Maria Jose Espona

ArgIQ

Resumen

La era de la información en la que vivimos, desde la explosión de Internet, presenta crecientes desafíos para la investigación académica en todos los campos del conocimiento. La gran disponibilidad de información nos genera complicaciones a la hora de seleccionar, en base a los objetivos planteados, el material a utilizar. Es aquí donde una disciplina surgida en el ámbito de la informática viene en nuestra ayuda.

La aplicación de la metodología de Calidad de Información, disciplina que comenzó a desarrollarse en el Instituto de Tecnología de Massachusetts (MIT) en la década del '90, nos ayuda a clasificar y seleccionar la información en base a una serie de atributos o dimensiones: precisión, objetividad, credibilidad, reputación, relevancia, con valor agregado, actualidad y oportunidad, completa, cantidad de información, entre otras. Este análisis multidimensional nos proporciona la posibilidad de realizar una evaluación más completa de los recursos disponibles.

En este trabajo describiremos esta disciplina, sus características y aspectos teóricos relevantes y se proveerán ejemplos de aplicación para verificar la utilidad del método.

Introducción

La disciplina de Calidad de Información lleva algo más de dos décadas entre nosotros, de crecimiento sostenido y ampliación constante de las áreas de aplicación de esta metodología. Aunque comenzó en el área de sistemas de información e informática, hoy es posible aplicarla a distintas áreas del conocimiento.

En primer lugar es fundamental aclarar antes de avanzar con los detalles de esta metodología la diferencia entre los conceptos *datos* e *información*, entendiéndose por *datos* el reflejo de hechos aislados despojados de sentido e *información*, como *datos* procesados que tienen sentido debido a las relaciones establecidas con otros *datos*. A veces nos perdemos entre papeles y textos y olvidamos dónde y dónde se encuentra la realidad. Lo único real es el hecho, al que nosotros usualmente no tenemos acceso y utilizamos ya sea *datos* o *información* procesados por otros, desconocidos en la mayoría de los casos, para tener alguna referencia al evento ocurrido.

Este punto es de desatacar, debido a que quienes transforman al hecho en dato y/o información lo hacen incluyendo en su proceso mental sus modelos mentales y a través de su percepción, con sus limitaciones y características propias (Figura 1). Estos matices, la mayoría de las veces desconocidos para los lectores, tiñen al hecho de subjetividades.

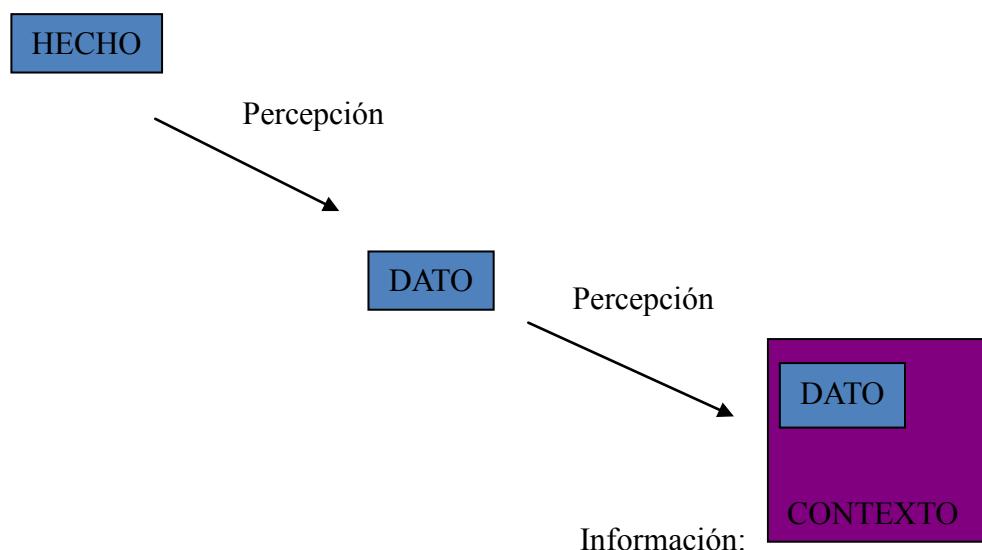


Figura 1: Relación entre hecho, dato e información. (Elaboración propia)

El objetivo de establecer una relación entre hecho, dato e información tiene que ver con que no podemos hablar de calidad de datos, sin saber de dónde vienen los datos o información sobre los cuales vamos a determinar o valorizar su calidad.

Asimismo, una vez que tenemos la información y la procesamos, obtenemos un producto de inteligencia que luego se transforma en conocimiento. Los datos son los ladrillos y pilares con los que construimos conocimiento y si ellos no son buenos, nuestro edificio se cae como un castillo de naipes. Si ahora consultáramos acerca de las cualidades que debería tener la información a académicos, políticos y gente en la calle, recibiríamos distintas respuestas, entre las cuales encontraríamos: exactitud, precisión y, en algunas circunstancias, oportuna, etc.

Pero, ¿es realmente así? Tomemos ejemplos cercanos:

1- ¿Es exacta, precisa u oportuna la información utilizada para el diseño de una política para de salud para combatir el dengue? Imaginemos un escenario donde un organismo brinde datos que, por un lado, indiquen como valor a 100.000 infectados, cuando en realidad son más y, por el otro, no de información alguna sobre su distribución geográfica.

2- ¿Fue exacta, precisa u oportuna la información que habría motivado la cancelación del despegue del Challenger? En realidad, la falla de O-rings ya era conocida.

Estos comentarios nos llevan ahora a más preguntas sobre la información y sus cualidades: ¿son suficientes la exactitud, precisión y la oportunidad? ¿Deberíamos considerar otros aspectos?

Los estudios realizados en el marco del Programa de Calidad de Información del MIT (Boston, EEUU) analizan este tema y plantean la existencia de al menos 15 dimensiones o aspectos destacados de la información que condicionan su calidad.

Cabe destacar que si bien el foco de la calidad de información pareciera estar solamente en el manejo de este insumo crítico, los procesos que ocurren en las organizaciones también la utilizan. Es decir, Calidad de Información no sólo se aplica al análisis de una problemática de la información, los datos y su calidad como elemento crítico de los procesos organizacionales, sino que también se ocupa del flujo de la información dentro de la empresa y de las decisiones tomadas.

La Calidad de Información y de Datos está recibiendo mayor atención por parte del gobierno de EEUU desde los ataques terroristas del 2001. Los medios de comunicación dieron a conocer reclamos que manifestaban que la mala información fue, al menos en parte, la responsable de la incapacidad de EEUU para prevenir los ataques o rastrear en tiempo real a los perpetradores (Fisher et al, 2006: vii).

Ejemplos de datos de pobre calidad abundan en nuestras vidas e incluyen a los ambientes industriales, gubernamentales, académicos, de salud y personales. Pierce dice: “la motivación de las organizaciones para entender y mejorar la calidad de datos y de información es más acuciante que nunca. Cada día es más frecuente que las organizaciones no mantengan contacto cara a cara con consumidores, vendedores, inspectores gubernamentales y aun con empleados” (Pierce, 2005: 265).

No hay duda de que vivimos en una era de información. El 93% de los documentos corporativos son creados electrónicamente. Cada año, billones de mails son enviados alrededor del mundo. Sino pensemos de qué manera nos comunicamos personal y profesionalmente y veremos que los mecanismos han cambiado muchísimo en los últimos años. Claramente, existe mucho más uso y dependencia de la información que antes, pero ¿qué ocurre si la calidad de la información es pobre?

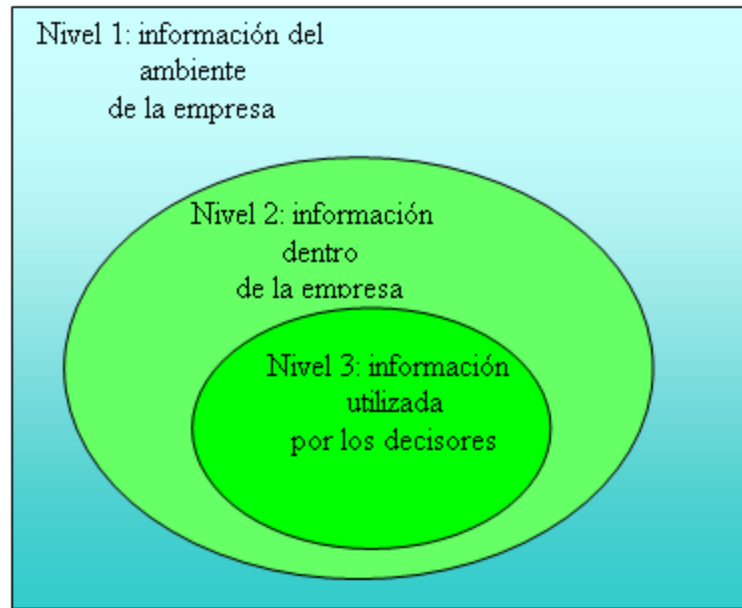


Figura 2: Trayectoria que realiza la información en su camino hacia el apoyo y la decisión (traducción de la autora de Pillkahn, 2008).

La cantidad de información disponible también es un problema, particularmente cuando ascendemos en la pirámide organizacional. A cada paso se realiza un recorte de la información provista, por ello es fundamental que sea el recorte correcto, de acuerdo a las necesidades del usuario

Es muy difícil determinar hasta qué punto una baja calidad de datos afecta al proceso de toma de decisiones y condiciona sus resultados (Chengalur-Smith, 1998). Sin embargo, sí se puede afirmar que con datos erróneos, las decisiones que se tomen basadas en ellos serán muy probablemente equivocadas. Cuando médicos, abogados, meteorólogos y mecánicos toman decisiones basadas en información de poca calidad, existe un gran riesgo de que sus conclusiones no sean correctas. Por otra parte, si los datos fueran 100% confiables, es mucho más probable que las conclusiones sean correctas (Fisher et al, 2006: 12). El problema es que muchas veces no sabemos o cuestionamos el nivel de calidad de la información con la que estamos trabajando y por ello vemos nuestro trabajo afectado sin saber muchas veces las razones.

En 1999, Ballou y Tayi plantearon que la mejor forma de definir calidad de datos es con la frase *fit for use*, es decir, apropiada para su uso. Los datos deben ser presentados en un formato que sirva a los propósitos del usuario, y deben estar expresados de forma tal que lo pueda comprender.

Sin embargo, *fit for use* es un término relativo y varía según sea el campo, el nivel dentro de la organización, la cultura organizacional y las preferencias personales. Lo que puede ser útil para los

ingenieros puede no serlo para los administradores. Las tablas que los ingenieros utilizaron para describir los problemas con los O-rings en el Challenger, hicieron poco para convencer a los ejecutivos de los peligros inherentes (Fisher y Kingma, 2001). A esto se lo denomina perspectiva del usuario y es fundamental que cuanto desarrollemos un trabajo, pensemos en quién lo va a usar, leer, es decir, no en cómo nos gusta a nosotros sino en cómo será mejor entendido.

Categorías y dimensiones

Pasando ahora a definir Calidad de Datos, podemos decir que, casi con seguridad, el grupo mayoritario considera que la calidad está en relación a la precisión y confiabilidad en los datos, mientras que el segundo grupo considera que está relacionado con la utilidad que posea en el contexto en la que la utilizarán. Sin embargo, podemos ir más allá y desarrollar un concepto multidimensional y plástico que les permita a los usuarios adaptar los criterios de calidad a cada situación en particular.

En este sentido, el Dr. Richard Wang, co-director del programa del MIT, junto con algunos colegas, determinaron este concepto multidimensional (Wang y Strong, 1996; Wand y Wang, 1996), basándose en la información recolectada de los usuarios y no en base a desarrollos teóricos y abstractos realizados detrás de un escritorio. Encontraron que no era fácil entender qué significaba calidad de datos para sus consumidores y entonces desarrollaron un marco de referencia que captura los aspectos de calidad que son más importantes para los consumidores de datos.

Los investigadores encontraron 179 atributos de Calidad de Información. Pero como eran demasiados para hacer un trabajo operativo, aplicaron distintas técnicas de análisis estadístico, para reducir el número a 15 y luego fueron agrupadas en 4 categorías, asignando una jerarquía y aportando un marco de referencia para trabajar con calidad de datos.

En este trabajo nos dedicaremos a las categorías utilizadas en el MIT. Sin embargo, otros autores e investigadores eligen las suyas. Esta es una de las muchas ventajas de este método: al trabajar con él podemos elegir cuáles son las dimensiones importantes para nuestro trabajo y definirlas en el propio contexto.

Las categorías de agrupamiento de dimensiones que presentamos son las siguientes:

Categoría	Definición
Intrínseca	Denota que la calidad del ítem está autocontenida, es decir, que el contexto no determina la calidad. Incluye las dimensiones: credibilidad, precisión, objetividad y reputación.
Contexto	Se tiene en cuenta al contexto como algo esencial. Incluye cinco dimensiones: valor agregado, relevancia, oportunidad, completa y cantidad de datos.
Representación	Refleja la importancia de la presentación de los datos y aspectos metodológicos. Incluye las dimensiones de interpretabilidad, facilidad de comprensión, consistencia representacional y representación concisa.
Accesibilidad	Lidia con la disponibilidad de datos y sobre cómo están protegidos del uso no autorizado. Las dimensiones son accesibilidad y seguridad.

Cuadro 1: Categorías de Calidad de Datos (elaboración propia)

En algunos casos, la calidad de los datos está directamente contenida en el dato mismo, por ejemplo, si es preciso o no. Cuando la calidad del dato es directamente reconocible a partir del dato, entonces se puede decir que es intrínseca a él. En otros casos, la calidad puede ser conocida sólo en el contexto de otros datos o por la manera en que el mismo es utilizado y aplicado. El dato puede ser preciso, pero si no es recibido a tiempo o es irrelevante, no le sirve al usuario para el propósito previsto. La tercera categoría, representación, está basada en la disponibilidad directa del dato. Algunos datos son precisos pero están codificados de tal forma que es muy difícil utilizarlos. Si éste no está presentado de una manera simple que represente algo con lo cual el usuario pueda trabajar fácilmente, entonces hay problemas representacionales. Por ejemplo, un conjunto de tablas con números, pero sin referencias a las unidades, puede tener significado para quien la hizo pero será de poca utilidad para otro lector.

La cuarta categoría es accesibilidad. Si el usuario que necesita los datos no los puede obtener o si las personas que no deben tener acceso a los datos lo tienen, entonces hay problemas de accesibilidad.

En detalle, el significado de cada una de las dimensiones es el siguiente:

Precisión: cuán exactamente es representado el mundo real. (Davenport, T.H., 1997 y Wang y Strong, 1996).

Credibilidad: el dato puede ser preciso, pero no es útil si el usuario no cree que lo sea.

Objetividad: si intervienen opiniones personales durante el proceso de creación del dato, puede ser que algunos individuos no confíen en él como si hubiera sido construido totalmente a través de medios objetivos. Asimismo, la objetividad influye en la dimensión de credibilidad.

Reputación: con el paso del tiempo, los datos construyen una reputación que los usuarios tienen en cuenta cuando evalúan la calidad de sus bases de datos e información. Más allá de la medida exacta de precisión, la reputación de la base puede disuadir al usuario de utilizarla en su proceso de toma de decisiones.

Valor agregado: extensión en la cual el dato es beneficioso y aporta ventajas a partir de su uso.

Relevancia: se refiere al grado en el cual el dato es apropiado y útil para una tarea determinada. Es decir, la aplicabilidad de un dato para un tema en particular por un usuario en especial (Tayi y Ballou, 1998; Orr, K., 1998; Redman, T.C., 1996).

Actualidad (y oportunidad): se refiere a la edad del dato (Wang y Strong, 1996), es decir, significa que el valor registrado no está desactualizado. Esta dimensión puede variar con el decisor y la circunstancia: un planificador estratégico puede utilizar información que tenga varios años, mientras que un gerente de producción necesita tener datos dentro de la hora.

Completa: se refiere al grado en el cual los datos están presentes en la colección (Ballou y Pazer, 1995). El Dr. Wang (1996) plantea que es “la extensión en la cual los datos se encuentran con la suficiente amplitud, profundidad y alcance para la tarea a realizar”. Con otras palabras, se enfoca en si todos los valores, para todas las variables, están registrados, almacenados y presentados.

Cantidad de datos: como su nombre lo indica, está especialmente relacionado con la sobrecarga de datos que puede ocurrir en un proceso y el impacto que posee debido a su tratamiento no óptimo.

Interpretabilidad: la información debe ser presentada en un lenguaje o unidades claras, lo cual incluye definiciones precisas de los términos.

Fácil de entender: si bien parece obvia, si no es planificada con un objetivo, puede no llegar a cumplirse. Esta dimensión establece que los datos deben ser claros, no ambiguos y fáciles de comprender (Wang y Strong, 1996).

Representación consistente: uso de formatos comunes de sistema en sistema y de aplicación en aplicación. Significa que los valores de los datos son los mismos en todos los casos (Ballou y Pazer, 1995; Ballou y Pazer, 2003). También implica que no hay redundancia en la base de datos. Profesionales de factores humanos han demostrado que esta dimensión es un factor clave en el diseño de sistemas (Shneiderman, 1986).

Representación concisa: en pocas palabras, sería corto y al punto.

Accesibilidad: extensión en la cual el dato está disponible o es fácil de obtener (Wang y Strong, 1996).

Seguridad de acceso: filtros para la obtención de los datos.

Categoría	Dimensión
Intrínseca	Precisión, Objetividad, Credibilidad, Reputación
De Contexto	Relevancia, con Valor agregado, Actualidad (y oportunidad), Completa, Cantidad de información
Presentación	Interpretabilidad (lenguaje y unidades), Fácil de entender, Representación concisa, Representación consistente (desde lo metodológico)
Accesibilidad	Accesibilidad, Seguridad de acceso (relacionado con oportunidad)

Cuadro 2: Categorías y Dimensiones (traducido de la autora de Wang et al. 2005)

Roles y relaciones

Otro aspecto muy importante en el método de Calidad de Información se refiere a los roles que asumen los individuos en las organizaciones y la comunicación que hay entre ellos.

Como investigadores, somos recopiladores de datos, sus custodios y los mismos serán utilizados por quienes lean nuestra publicación. Pero también podemos producir un informe solicitado para que sea tomada una decisión, por ejemplo, una encuesta sobre usos de los medios de transporte (propio, público o privado), con el objetivo de mejorar la red de transporte en la ciudad.

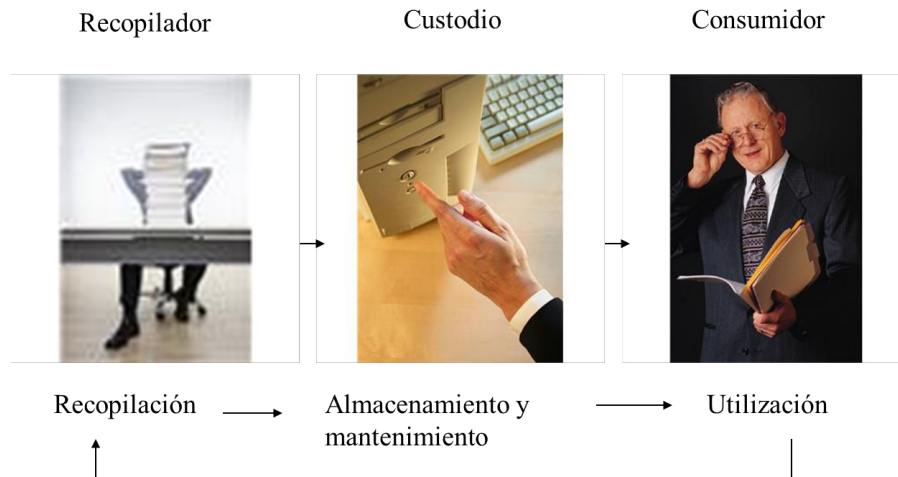


Figura 3: roles en Calidad de Datos

Dado que la calidad de los datos es definida por el usuario o consumidor (*fit for use*) la retroalimentación entre el recopilador, el custodio y el consumidor es fundamental para así asegurar que se cumplen los requisitos de calidad previstos.

La comunicación efectiva de las necesidades del usuario por parte de él, así como también la comprensión por parte del recopilador son fundamentales y nos remiten a los elementos de comunicación que nos enseñaron en la escuela primaria.

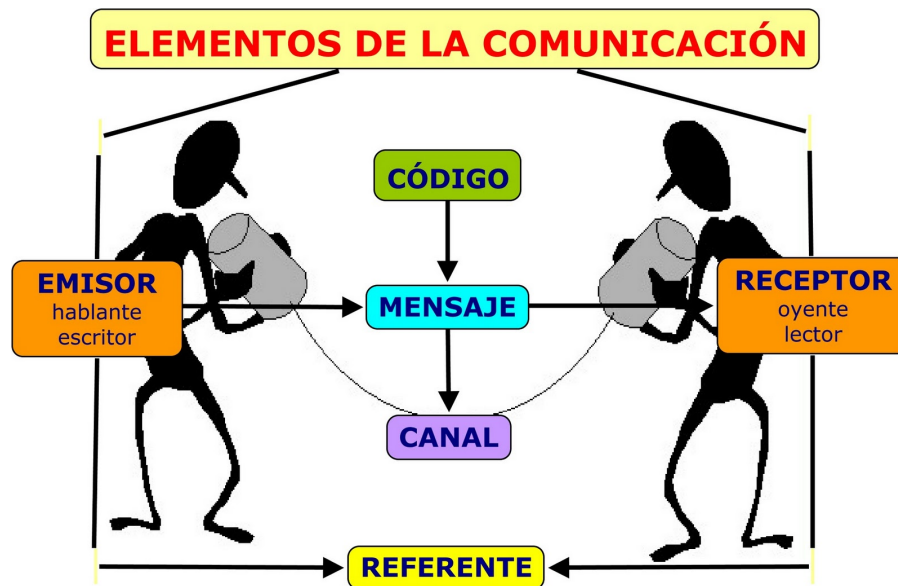


Figura 4: elementos de la comunicación¹

¹ Imagen extraída de

https://www.google.com.ar/search?q=comunicacion&newwindow=1&source=lnms&tbm=isch&sa=X&ei=Q6g9VLjKDceONrDpgdAM&ved=0CAgQAUoAQ&biw=1920&bih=979#facrc=&imgdii=&imgsrc=ju2m99QXn9Vi2M%253A%3BWDu_A2ZISXfJKM%3Bhttp%253A%252F%252Ftucajadetalentos.files.wordpress.com%252F2014%252F04%252Fcomunica.jpg%3Bhttp

Ciclo de gerenciamiento de calidad total de datos

Finalmente, el último elemento de este modelo tiene que ver con una visión sistémica de la calidad de datos, que aúna los dos aspectos antes mencionados, categorías y dimensiones y roles y relaciones. Cabe señalar que, cronológicamente, es el primero en ser aplicado, dado que primero identificamos al objeto de estudio a ser medido en cuanto a su calidad, aplicando en ese estadio las dimensiones y su valorización, para luego ver la construcción de roles.

Las etapas del ciclo son:

Definir: establecer criterios de calidad de datos. Concepto multidimensional (va más allá de precisión)

Medir: calidad de datos con o sin herramientas informáticas

Analizar: con modelos, métodos y principios de Calidad de Datos

Mejorar: mediante la implementación de las medidas necesarias

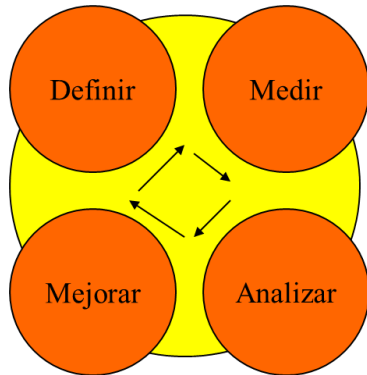


Figura 5: Ciclo de calidad total de datos

También puede aplicarse este ciclo a la evaluación de procesos de investigación, como por ejemplo, evaluación de encuestas, relevamiento de datos, etc.

Caso de estudio: Agentes biológicos en América del Sur

Si consideramos que los agentes de guerra biológica son sólo patógenos presentes en la naturaleza que fueron incluidos en programas de armas (Guillemin 2004: 2), vemos que la guerra biológica no es algo tan ajeno a nuestra realidad.

Para conocer si existe o no un ataque biológico en la región debemos conocer qué enfermedades se presentan normalmente en el área y cuáles son sus características. Pensando en esta idea es que

intentamos, a partir de la lista virus del Grupo Australia², determinar cuál era la situación epidemiológica en la región, la década pasada.

Gran sorpresa nos llevamos cuando vimos que los datos eran totalmente dispares en todos los países y para todas las enfermedades.

Pero antes de abandonar el objetivo previsto dados los problemas que presentaban la información disponible en cuanto al grado de detalle con el que podíamos trabajar, le dimos otro enfoque y buscamos determinar al menos la presencia de virus listados de identificar problemas con los datos que pudimos acceder.

Para ello, explotamos las fuentes públicas disponibles, utilizando a Google como motor de búsqueda y también los motores de búsqueda de cada página web perteneciente a las organizaciones consideradas, como por ejemplo la OMS, la OPS, CDC. Para nuestro análisis, consideramos como marco temporal diez años a partir de 1998 (es decir, hasta 2008) y, asimismo, registramos la fuente de la información más relevante (desde el nivel supranacional hasta el local) a fin de calificarla y categorizarla, usando la metodología de Calidad de Información (Lee et al. 2006, Fisher et al. 2006, Wang & Lee 2005).

Agente de Guerra Biológica	País
Virus de la fiebre del Dengue	Argentina, Bolivia, Brasil, Chile, Colombia, Ecuador, Paraguay, Perú.
Virus de la Encefalitis equina del este	Argentina, Brasil
Hantavirus	Argentina, Bolivia, Brasil, Chile, Paraguay, Uruguay, Venezuela
Virus Junín	Argentina
Virus de la coriomeningitis linfocítica	Argentina, Bolivia, Brasil, Chile, Colombia, Ecuador, Paraguay, Perú, Uruguay, Venezuela
Virus Machupo	Bolivia

² http://australiagroup.net/en/human_animal_pathogens.html, último acceso, 14/10/2014

Encefalitis equina venezolana	Argentina, Brasil, Colombia, Ecuador, Perú, Venezuela
Encefalitis equina del oeste	Argentina
Virus de la fiebre amarilla	Argentina, Bolivia, Brasil, Colombia, Ecuador, Paraguay, Perú, Venezuela
Virus Oropuche	Argentina, Brasil, Perú
Virus Rocio	Argentina, Brasil
Virus de St. Louis	Argentina, Bolivia, Brasil, Chile, Colombia, Ecuador, Paraguay, Perú, Uruguay, Venezuela
Virus Sabia	Brasil
Flexal Virus	Brasil
Virus Guanarito	Venezuela
Virus del Nilo Occidental	Argentina

Cuadro 3: Agentes de Guerra Biológica presentes en América del Sur.

Es importante señalar que la información mostrada en el cuadro 3 fue construida teniendo en cuenta más de una fuente de información para cada virus, durante el marco temporal señalado y teniendo al indicador más pequeño de presencia como una respuesta afirmativa. El último concepto se basa en las complicaciones en determinar la presencia de una enfermedad específica, debido a la situación geográfica, las características de los laboratorios involucrados, la política de atención de salud, los reglamentos de control, políticas de comunicación de enfermedades y las dificultades inherentes a las enfermedades y los agentes etiológicos bajo consideración.

Un análisis de calidad de la información epidemiológica recopilada nos dio los siguientes resultados:

1. Categoría: Intrínseca

Cuando se trata de información epidemiológica, credibilidad y reputación vienen juntas, e incluso las observamos como cercanamente relacionadas a la objetividad. Incluso cuando la precisión suena más importante, una información cualitativa de una fuente creíble puede ser más útil que un dato

cuantitativo de una fuente no creíble. Por lo tanto, si queremos conocer el panorama general de la situación epidemiológica de una región o país, necesitamos información que sea creíble, objetiva y con una buena reputación. Así, un dato cualitativo será suficiente. Pero si somos decisores diseñando una estrategia de salud, necesitaremos también información precisa con datos cuantitativos.

2. Categoría: Contextual

En la situación en la que estamos recopilando información epidemiológica, encontramos toneladas de la ella provenientes de diferentes fuentes (medios de comunicación, ONGs, Organizaciones Supranacionales, publicaciones científicas, noticias de Ministerios de Salud, etc.), ciertos datos duplicados y otros datos únicos encontrados escondidos en artículos no relacionados.

Un tema crítico es la cantidad de datos, aspecto difícil de tratar. Dependerá de la cantidad de tiempo y recursos del investigador si éste es o no un problema. De cualquier modo, Internet nos muestra una enorme cantidad de información, en la mayor parte de los casos imposibles de procesar manualmente.

Por otro lado está lo exhaustivo de la información, ya que, como todos sabemos, tener mucha información no significa que uno tenga toda la información que necesite.

Otro aspecto es el valor agregado (¿cuánta información está duplicada o no hace una nueva contribución?) y la relevancia (¿es la nueva información importante, teniendo en cuenta los objetivos de la investigación?). Sabemos que ambas situaciones parecieran ser un problema de poca dimensión, pero es necesario leer todos los documentos para saber si la nueva información añade valor y si es relevante.

La oportunidad se relaciona con la edad de los datos y si los mismos están actualizados o no. En el campo de la salud, éste es un aspecto crítico; tener información actualizada hace una gran diferencia, tanto en la planificación estratégica de la salud, como en el abastecimiento de medicamentos en una farmacia.

3. Categoría: Representación

La información epidemiológica debe ser de fácil entendimiento e interpretación: número de personas enfermas, número de muertes, área abierta por el vector de la enfermedad, estrategia de transmisión, etc. Por ejemplo, si se tratase de cantidad de casos, debería quedar claro si dicho número incluye únicamente a las personas enfermas o también a aquellos muertos por la enfermedad (presentación concisa y representación consistente).

La consistencia representacional dentro de una organización nos permitirá, en nuestro caso de estudio, seguir la evolución de una enfermedad y confiar en los resultados.

4. Categoría: Accesibilidad

El equilibrio entre el acceso y la seguridad de acceso es posible observarlo en la mayoría de los campos de investigación. En el caso de los estudios epidemiológicos, es importante que el acceso a la mayor cantidad de información posible esté abierto para que los datos puedan ser analizados por expertos, mientras que esté restringida a actores estatales o a organizaciones terroristas en búsqueda de realizar un ataque con armas biológicas.

Conclusiones

Esta metodología de trabajo avanza continuamente en todos los campos del conocimiento, pasando de ser una herramienta para la toma de decisión empresarial, a una de aplicación en investigaciones de distintos tipos.

La plasticidad que aporta para la selección de criterios de calidad la hace apta para distintas disciplinas de las ciencias sociales y naturales.

Las aplicaciones realizadas hasta ahora en epidemiología de enfermedades y determinación de límites geográficos han dado muy buenos resultados.

Finalmente, la accesibilidad y disponibilidad de esta herramienta para el mundo académico y empresarial la transforma en una opción muy interesante, en estos días de sobreabundancia de información.

Bibliografía

- BALLOU, D. & G. TAYI (1999). Enhancing Data Quality in Data Warehouse Environments. Communications of the ACM: 42 (1), pp. 73-78.
- BALLOU, D.P. & H. PAZER (2003). Modeling Completeness vs Consistency in Information Decision Contexts. IEEE Transactions on Knowledge and Data Engineering: 15(1).
- BALLOU, D.P. & H.L. PAZER (1995). Designing Information Systems to Optimize the Accuracy-Timeliness Tradeoff. Information Systems Research: 6(1), pp. 51-72.
- BELL, T.E., & ESCH, Karl (1987). The Fatal Flaw in Flight 51-L. IEEE Spectrum: 24(2), pp. 36-51.

- CHENGALUR-SMITH, I. & H. PAZER (1998). Decision Complacency, Consensus and Consistency in the Presence of Data Quality Information. en Conference on Information Quality. Cambridge, MA.
- CLUNAN, ANNE L., PETER R. LAVOY & SUSAN B. MARTIN (2008). Terrorism, War or Disease? Unraveling the use of Biological Weapons. Stanford: Stanford University Press.
- DAVENPORT, T.H. (1997). Information Ecology. New York, NY: Oxford University Press.
- FISHER, C., E. LAURIA, I. CHENGALUR-SMITH & RICHARD Y. WANG (2006) Introduction to Information Quality. MITIQ: Cambridge.
- FISHER, C.W. & B.R. KINGMA, (2001). Criticality of Data Quality as Exemplified in Two Disasters. Information & Management: 39(2), pp. 109–116.
- HEUER, RICHARDS J. JR., (1999). Psychology of Intelligence Analysis. Center for the Study of Intelligence: Washington.
- LYMAN, PETER Y HAL R. VARIAN (2003). How much information? http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/printable_report.pdf (último acceso el 30/11/09).
- GUILLEMIN, J. (2004). Biological Weapons: From the Invention of State-sponsored Programs to Contemporary Bioterrorism. New York: Columbia University Press
- PILLKAHN, U. (2008). Using Trends and Scenarios Tools for Strategy Development: shapping the future of your enterprise. Berlin and Munich: Siemens.
- OMB, Information Quality Act, in Section 515. 2001. Office of Management and Budget; The Excecutive Office of the President, 17.
- ORR, K. (1998). Data Quality and Systems Theory. Communications of the ACM: 41(2), pp. 66–71.
- PIERCE, E.M. (2005). Introduction to Information Quality, en Information Quality, R.Y. Wang, E.M. Pierce, S.E. Madnick & C.W. Fisher., (eds.). M. E. Sharpe: Armonk.
- PILLKAHN, U. (2008). Using Trends and Scenarios Tools for Strategy Development: shapping the future of your enterprise. Berlin y Munich: Siemens.
- RECCE, J. (2009) Comunicación personal.
- REDMAN, T.C. (1996). Data Quality for the Information Age. Norwood, MA: Artech House, Inc.
- ROBERTS, N.C. (1992). Reconstructing Combat Decisions: Reflections on the Shootdown of Flight 655. Monterey: Naval Postgraduate School.

- ROGERS, W., & ROGERS, S. (1992). Storm Center The USS Vincennes and Iran Air Flight 655. Annapolis, MD: Naval Institute Press.
- ROGERS, W.P. (1986) Report of the Presidential Commission on the Space Shuttle Challenger Accident. Washington, D.C.: United States Government Printing Office.
- SHNEIDERMAN, B. (1986). Designing Menu Selection Systems. Journal of the American Society for Information Science: 37(2), pp. 57–70.
- TAYL, G. & D.P. BALLOU (1998). Examining Data Quality. Communications of the ACM: 41(2), pp. 54–57.
- WAND, Y. & R.Y. WANG (1996). Anchoring Data Quality Dimensions in Ontological Foundations. Communications of the ACM: 39(11), pp. 86–95.
- WANG, R & Y. LEE. Apuntes del curso: Information Quality: Principles and Foundations [ESD.IQ1] March 21-25, 2005, MIT, Cambridge.
- WANG, R.Y. & D. STRONG (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems: 12(4), pp. 5–34.